

Weekly Report

1 Done

1.1 Vast Presentation

I took most time to do the slides this week. In next two weeks, I will prepare for presentation.

1.2 Discussion with Prof. Ren

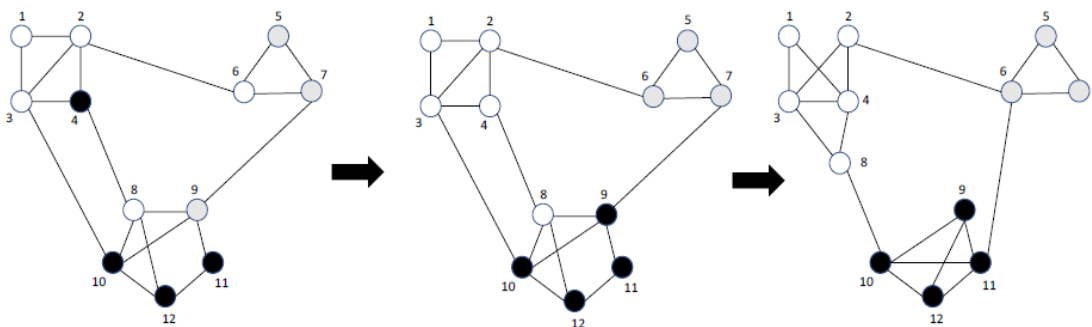
- He learnt about the usual contributions in our field.
- In their field (security), they emphasize on adversary models and prefer quantitative optimization goals. Prof. Ren said that it is meaningful to identify the similarity of patterns of original data and processed data. However, it is challenging to us as well. As for now, we can only employ simple metrics. In his opinion, automatic algorithms should tell users the best solutions based on user-defined optimization goals (utility needs and privacy needs). In my opinion, it is still necessary to show users the data to define appropriate goals. Besides, he agreed with the opinion that users need the explanation of the privacy preserving processes.
- We can work together to do something that makes sense in both fields.
- Zhan Qin, one of his students, has published a paper on differential privacy for graph data.

1.3 Paper Reading

- *Generating Synthetic Decentralized Social Graphs with Local*

Differential Privacy (Zhan Qin's paper)

This work proposes LDPGen, a novel multi-phase technique that incrementally clusters users based on their connections to different partitions of the whole population.



The approach takes three phases as shown above.

- Optimize group number and refine partition.
- Further refine partition with optimized group number.

- Generate synthetic graph based on refined user partition.

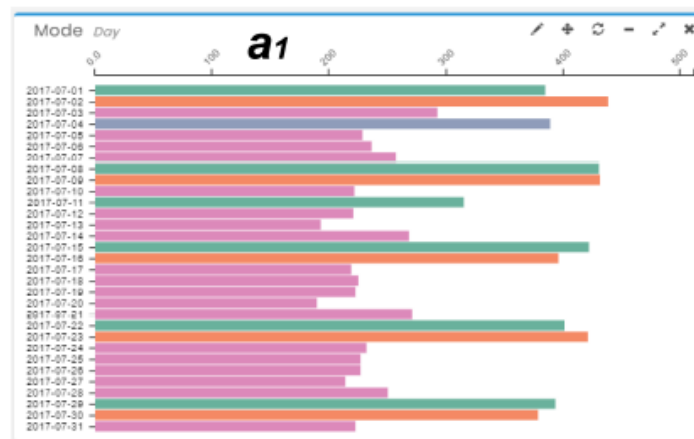
The first two phases are implemented to group nodes. Our idea is grouping nodes by community. Both of them need grouping nodes. The common points can be extracted as a pipeline. Moreover, we can compare the two approaches.

- TPFlow: Progressive Partition and Multidimensional Pattern

Extraction for Large-Scale Spatio-Temporal Data Analysis (VAST Best Paper)

The main contribution of this work is a novel piecewise rank-one tensor decomposition algorithm which automatically seeks for and recommends the best way to slice multidimensional ST data and extract the multidimensional trend for comparison and visual summarization.

In their approach, the patterns among all slices are compared. Therefore, they can observe the periodic patterns like weekend traffics.

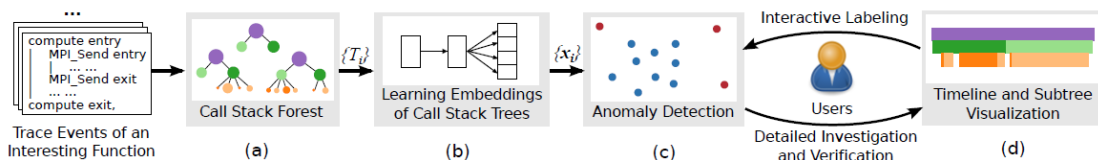


Recommendation plays a more and more important role in visual analytics. We can attempt to provide this feature in our project.

- A Visual Analytics Framework for the Detection of Anomalous Call

Stack Trees in High Performance Computing Applications (VAST HM)

This work is about Anomaly detection.



They first generate CSTrees from the trace events. Next, they construct feature vectors using stack2vec. The candidate anomalous CSTrees are detected in the forest. Then, users can investigate the candidates in detailed visualization. Labeling information provided by the user will be fed back to update the anomaly detection model.

- Formalizing Visualization Design Knowledge as Constraints:

Actionable and Extensible Models in Draco (InfoVis Best Paper)

In this paper, authors propose a collection of constraints, in conjunction with a method to learn weights for soft constraints from experimental data.

- Mapping visualization specifications to logical facts.
- Representing design knowledge as constraints.
- Completing specifications by solving constraints.

They also implement their approach in Draco, a constraint-based system based on Answer Set Programming (ASP).

2 Work Hours

I went to the medical examination on Monday, 8 am-9:20 am. On Saturday afternoon (about 1:30-4:00), I went to Yuquan campus to discuss with Prof. Ren. In the rest of weekdays, I worked during 9:00~11:30, 13:30~5:30 and 19:00~20:30. On weekends (Thursday and Friday in this week), I read paper in my dormitory.

3 Progress

Item	Deadline	Current progress	Remark
Vis presentation	10.24	Slides are done.	
Go abroad	11.18	Got J1 visa.	
Privacy program	10.31	Implementing existed approaches.	